Selection, Firm Turnover, and Productivity Growth: Do Emerging Cities Speed up the Process?

Patricia Jones
University of Oxford

Taye Mengistae
World Bank

Albert Zeufack
World Bank

February 2018

This paper is a part of a Global Research Program on Spatial Development of Cities, funded by the Multi Donor Trust Fund on Sustainable Urbanization of the World Bank and supported by the UK Department for International Development. We wish to thank Dewi Spikerman for excellent research assistance.

Abstract

In this paper, we identify and estimate the impact of firm entry and exit on plant-level productivity in Ethiopia as part of a selection mechanism that might be driving aggregate productivity growth in cities. Specifically, we are interested in how firms' entry and exit contribute to the pace of factor reallocation and TFP growth within industries—and whether these processes occur in higher numbers in larger cities. We carry out this analysis using establishment-level census data from Ethiopia which cover the period 2000 to 2010. Importantly, these data include information on plants' physical outputs and their prices which allow us to distinguish between revenue-based measures of total factor productivity (TFPR) and those based on physical productivity (TFPQ). Our analysis reveals that these two measures generate very different results under imperfect competition, suggesting that physical productivity measures (TFPQ) are better suited to examining firm dynamics when local producers have some degree of market power. In addition, we find that less productive (higher cost) firms are more likely to exit than their more productive (lower cost) rivals—but only when we control for producers' transport costs. This is consistent with the probability of firm exit being higher when transport costs are lower.

Key words: Productivity, Pricing and Market Structure, Urbanization
JEL Codes: D24, L11, R11

I. Introduction

Africa is urbanizing fast. Currently, 472 million people live in urban areas across the continent and this number is expected to double in the next twenty-five years (United Nations, 2015). Africa's pattern of urbanization, however, is different from that of other developing regions. Elsewhere, increased urbanization has been accompanied by a rise in the share of manufacturing in economic activity. Globally, there is strong correlation between urbanization and the expansion of manufacturing. For most countries, the manufacturing share rises with urbanization until about 60% of the population lives in cities and manufacturing accounts for about 15% of GDP. By contrast, the relationship between urbanization and manufacturing in Sub-Saharan Africa is relatively flat (Figure 1). This unique pattern of growth has been described as "urbanization without industrialization" (Fay and Opal, 2000; Jedwab, 2013; Gollin et al, 2016).

The development literature has highlighted two distinct forms of structural change that countries can follow as they urbanize. The first path involves the typical movement of workers out of agriculture and into manufacturing. This type of structural change results in the growth of "production" cities in which tradable goods are produced for both domestic and international markets. This is the path which has been taken by most countries in Europe, Latin America, and Asia (Bairoch, 1988).

The second path is meant to reflect the recent experience of several African countries which have large, natural resource endowments. In these countries, positive productivity shocks to the resource sector have shifted workers out of the food and tradable sectors and primarily into the non-tradable sector. The surplus income generated from these productivity shocks has caused a disproportionate rise in the demand for urban goods and services relative to food. This additional demand has been met largely through imports except in the case of urban services which is provided by the local labor force. The net result is a rise in the level of urbanization and an increase in the share of employment in the urban, non-tradable sector. This type of urbanization is driven mostly by consumption rather than production, resulting in the emergence of "consumption" cities (Gollin et al, 2016).

In this paper, we study the link between urbanization and industrialization in Africa from a different perspective. Our focus is on the role of competition-driven selection mechanisms in raising aggregate productivity. Specifically, we are interested in two mechanisms: 1) the contribution of firms' entry and exit to factor reallocation and its impact on total factor productivity (TFP) growth; and 2) the importance of cities in speeding up this process. One hypothesis about cities it that they lead to tougher competition in urban markets. They do this by attracting new firms into urban areas that, given a fixed amount of land, reduces the physical distance between firms (a lá Salop, 1979). Closer producer spacing makes it harder for less productive (or higher cost) firms to compete against their more productive (or lower cost) competitors, resulting in less productive firms exiting the market and more productive firms entering the market. Simultaneous entry and exit of firm—aka firm turnover—can thus be an important driver of factor reallocation and productivity growth. These propositions are formalized in a general equilibrium model of monopolistic competition laid out in Melitz and Ottaviano (2008), which this paper draws on.

To our knowledge, this is the first study which investigates whether the urbanization process in a developing country is linked to increased factor reallocation through higher firm turnover. If larger cities increase the pace of factor reallocation, as is the implication of the Meltiz-Ottaviano model, Africa's growing cities may boost its economy-wide productivity regardless of whether urbanization leads to agglomeration economies. Indeed, market selection via firm entry and exit has been found to explain a large proportion of the variation in manufacturing productivity in other countries. For example, Baily, Hulten, and Campbell (1992) estimate that 50% of the productivity growth in US manufacturing during the 1970s and 1980s is attributed to factor reallocation. At the aggregate level, much of the variation in TFP growth across countries is explained by factor reallocation within narrowly (usually 4-digit ISIC) industries (see Foster Haltiwanger, and Syverson (2008); Syverson (2011); and Bartlesman, Haltiwanger, and Scarpetta, (2013)). Identifying the underlying determinants of TFP growth is therefore important not only for understanding firm dynamics within countries but also because TFP gaps are major sources of cross-country differences in per capita income (Prescott, 1998; Hall and Jones, 1999; Restuccia and Rogerson, 2008, 2013).

What drives the links between urbanization and increased factor reallocation? An implication of the Melitz-Ottaviano model is that high transport costs can reduce the scale of market selection via firm entry and exit. This is partly because, as pointed out by Syverson (2004a, 2004b), transport costs affect the ease with which consumers substitute the output of one producer for that of another. If high transport costs lower product substitutability, less productive firms can survive, even in long-run equilibrium. An important policy question following from this is how much of a constraint urban transport costs pose to African firms. Are high transport costs (or other factors that increase barriers to substitution between varieties of a product) reducing the pace of factor reallocation in the region?

We investigate these questions using a 10-year panel of manufacturing firms from Ethiopia.[1] This panel contains annual production data for the period 2000 to 2010 and covers all manufacturing firms employing 10 or more employees.[2] It is arguably the most comprehensive longitudinal dataset on manufacturing firms in Sub-Saharan Africa at this moment. Industries are classified in the dataset at the 4-digit ISIC level and each producer is geo-referenced at the town level. Importantly, the Ethiopian data include producers' physical outputs, q_i, along with their respective prices, p_i. This allows us to distinguish between revenue-based measures of total factor productivity (TFPR) and those based on physical outputs (TFPQ). As is standard in the literature, we define TFPR as the value of revenue $(p_i q_i)$ per input unit (x_i) and TFPQ as the number of physical units produced per unit of output (q_i/x_i). TFPQ measures the technical efficiency of a plant.

As discussed by Foster, Haltiwanger, and Foster (2008), it is important to distinguish between a plants' TRPR and TFPQ when estimating the effects of plant turnover on an industry's productivity. If all firms are price takers, simultaneous entry and exit processes leads to a selection outcome in which the industry's least productive firms (in the sense of having have lower TFPQ) exit the market in (long-run) equilibrium. However, it is also possible to have an alternative scenario in which a plant's revenue productivity (TFPR) is a better predictor of its

[1] This panel is compiled from ten years of the Ethiopian Survey of Large and Medium Scale Manufacturing Industries which, despite its name, is a census of all manufacturing firms with 10+ employees.
[2] Data from 2005 are dropped because a survey was conducted during that year rather than a census.

survival than its physical productivity (TFPQ). Such a case could arise if some plants in the industry exercise a degree of market power that allows them to charge higher prices for their product than others. If this were the case, plants with higher TFPR could survive in the long run, even if they were less productive (in the sense of having lower TFPQ) than plants exiting the market. In such cases, empirical studies that measure establishment-level productivity using TFPR might thus overestimate the "true" link between a firm's productivity and the probability of its survival.

This kind of measurement error arises when factors other than inter-firm gaps in factor productivity determine inter-firm price differentials. Such factors include the many sources of product differentiation that firms use to lower product substitutability as well as idiosyncratic demand differences that can arise in markets. While such factors drive a wedge between marginal factor productivities and product prices in monopolistic markets, they are not easy to observe or measure. The complicates the empirical task of identifying and evaluating firm turnover, market size, and transport costs as observable and measurable elements of spatial competition.

To get around this problem, we control for the influence of product differentiation simply by focusing on Ethiopian industries that produce only relatively homogeneous goods. That is, industries where vertical product differentiation is not likely to be a significant source of product market power. Our goal is to ensure that the price differentials that we observe reflect horizontal product differentiation (consumer preferences over products driven by supplier locations) rather than vertical product differentiation (consumer preferences over products driven by quality differences). We focus on the following 4-digit industries: non-rice flour (ISIC=1531), white pan bread (ISIC=1541), and cinder blocks (ISIC=2695). To the best of our judgement, these products represent the most homogeneous of goods we can identify in the dataset. In total, our pooled sample covers more than 2,500 plant-year observations.

Our analysis reveals that, in Ethiopia, low-productivity plants (in terms of TFPQ) are more likely to exit the market than high-productivity plants but only when we control for producers' transport costs. It turns out that the same selection mechanism that gives local producers some degree of market power is weakened by high transport costs. Our finding is consistent with both the Melitz-Ottaviano model and Syverson's (2004a) model which shows how high transport costs can lower spatial product substitutability, making it easier for less productive plants to survive in long-run equilibrium. We also find weak evidence of a market size effect on the same selection mechanism in as far as the inverse correlation between productivity and exit probability is stronger in Ethiopia's primate city, Addis Ababa, than in secondary cities.

The rest of the paper is organized into three sections. Section II presents the theoretical framework of the model of monopolistic competition laid out in Melitz and Ottaviano (2008) and the key hypotheses that we test empirically. Section III discusses our estimation strategy and presents our main results. Our estimation strategy draws heavily on Foster, Haltiwanger and Syverson (2008). Section IV concludes the paper.

II. Theoretical Framework: Demand side and supply side drivers of market selection and productivity growth

The Melitz-Ottaviano model belongs to a class of models of monopolistic competition in which firm dynamics and aggregate productivity growth are driven by a process of market selection whereby producers below a productivity threshold are forced to exit the market, ceding market share to more productive and expanding entrants and survivors. Briefly, it is a model in which aggregate, industry-wide productivity and the scale of entry and exit all depend on three demand side variables: 1) the size of the industry's product market; 2) the degree of product differentiation within the industry; and 3) the cost of transport to the point of delivery. The effect of supply side factors on productivity is transmitted via a fixed sunk cost of entry that is assumed to be exogenous to the model.

A. Product differentiation and market size as demand side factors in competition

In the model an industry consists of a continuum of N producers, indexed by $i : i \in \Omega$, producing distinct varieties of a product to meet demand from a continuum of L consumers who are assumed to have identical preferences over the varieties per the utility function

$$U = q_o + \alpha \int_{i\in\Omega} q_i di - \frac{1}{2}\gamma \int_{i\in\Omega} q_i^2 di - \frac{1}{2}\eta\left(\int_{i\in\Omega} q_i d_i\right)^2 \tag{1}$$

where q_i represents consumption of the output of $i : i \in \Omega$; $q_0 : q_0 > 0$ is quantity consumed of a unique numeraire good; α and η, are constants measuring the ease of substitution between the numeraire and varieties of the differentiated product; and γ is the degree of product differentiation and thus is an inverse measure of the ease of substitution among varieties.

For all varieties production involves the use of inelastically supplied homogenous labor as the only factor input to produce a differentiated good at a constant marginal cost, c, excluding transport costs. Both production and consumption take place in a multiplicity of locations, which could be cities, regions, or even countries. It is assumed that at least some of the produce of each location is consumed locally but varieties are imported from other locations subject to transport costs.

Without loss of generality we consider the simplest case whereby all economic activity takes place in just two distinct locations (or cities), h and l, such that l is the larger of the two local markets in the sense that $L^l > L^h$. Everyone is assumed to consume positive quantities of the numeraire in utility function (1) at unit price $p_i^l \equiv p_{iD}^l(c) + p_{iX}^l(c) = \alpha + \gamma q_i^l - \eta Q^l$, where D and X index, respectively, local sales and exports to the other location, and $Q^l = \int_{i\in\Omega} q_i^l d_i$. The unit price is assumed to be consistent with the aggregate inverse demand function for each variety that can be inverted into a demand system across varieties as

$$q_i^l = \frac{\alpha L^l}{\eta N^l + \gamma} - \frac{L^l}{\gamma} p_i^l + \frac{\eta N^l}{\eta N^l + \gamma}\frac{L^l}{\gamma}\bar{p}^l \qquad \forall i \in \Omega_l^* \tag{2}$$

where N^l is the number of varieties sold in location l (equal to the number of firms selling in that location including both local producers and exporters) based in h, L^l is the number of consumers in the same location, $\bar{p}^l = \frac{1}{N^l}\int_{i\in\Omega_l^*} p_i^l di$ is the average price in the location where

$\Omega_l^{\ *} \subset \Omega_l$, $\Omega \equiv \Omega_l \cup \Omega_h$ and $p_i^l \leq p_{Max}^l$, where p_{Max}^l is the price ceiling that would reduce the demand for any variety to zero in that location. The price ceiling in the location is given by

$$p_{Max}^l = \frac{1}{\eta N^l + \gamma}\left(\gamma\alpha + \eta N^l \bar{p}^l\right) \tag{3}$$

A precise indicator of the extent of competition that producers face in the product market is the elasticity of demand. For any variety, i , this is given by $\varepsilon = \left[\left(\frac{p_{Max}^l}{p_i^l}\right) - 1\right]^{-1}$ because of equation (3). Sellers in either location are said to face greater product market competition: (a) the greater is the aggregate demand for the industry's output as indicated by α ; (b) the lower is the industry average price, $\bar{p}^l$; (c) the greater is the ease of substitution between varieties, that is, the smaller is γ ; and (d) the larger is the number of sellers, N^l . The last result follows from the fact that, other things being equal, the price elasticity of demand is higher the larger is N^l because an increase in the number of sellers reduces the price ceiling, p_{Max}^l , in a location. Demand is also more price elastic for the more expensive varieties other things being equal.

B. Transport cost as a supply side factor in competition and productivity

While the same differentiated good is (produced and) delivered locally at the same unit cost, c , in either location, it is assumed that the delivery of identical output to the other location entails additional transport costs, $\tau^l c$ such that $\tau^l > 1$. This implies that the product market is segmented between the two locations by positive transport costs so that each producer maximizes its profits from local sales independently of its profits from exports to the other location.

Transport costs make it more difficult for firms to sell outside of the local market in that they need to charge a higher breakeven price than they do selling locally. Let c_D^l be the highest of the unit costs for profitably delivering any variety in location l while c_X^l is the highest of the unit costs of profitable shipments of the same variety to the other location. By assumption c_D^l is the same as the unit cost of the marginal (or highest -cost) local supplier in location l. That is, the producer that has the highest cost among the firms selling locally in l and is consequently just breaking even by charging the highest of the local prices observed, p^l_{Max}. We thus have $c^l_D = \sup\{c : \pi^l_D(c) > 0\} = p^l_{Max}$, where $\pi^l_D(c)$ is maximized profits from local sales. Also c^l_X is the highest possible marginal cost of shipment of a variety to the other location, h, in the sense that $c^l_X = \sup\{c : \pi^l_X(c) > 0\} = \frac{p^h_{Max}}{\tau^h}$, where $\pi^l_X(c)$ is maximized profits from exports to the other location . But this implies that $c^h_X = \frac{c^l_D}{\tau^l}$, which means that no producer can just breakeven by making any shipment of its output to the other location without charging a higher price than it would charge if it were selling the same output locally. Moreover, given any unit price of a variety, more of the variety is sold locally at that price than would be shipped to the other location.

C. Sunk entry costs as a supply side factor in competition and productivity

Let $q^l_D(c)$ be the quantity that a firm based in location l sells locally at the profit maximizing unit price, $p^l_D(c)$, and let $q^l_X(c)$ be the quantity of its shipment to the other location at the profit maximizing unit price, $p^l_X(c)$. Maximized profits from local sales and exports to the other location are thus given respectively by $\pi^l_D(c) = [p^l_D(c) - c]q^l_D(c)$ and $\pi^l_X(c) = [p^l_X(c) - \tau^h c]q^l_X(c)$

Firms make the decision on whether to produce only after having incurred a fixed sunk cost of entry, f_E, that is assumed to be invariant between locations. This decision is based on each firm's assessment of the profits that it expects to make by supplying either or both markets. The expected profits in turn depend on the firms' draw from the cost distribution, $G(c)$, across all potential producers. Given $G(c)$ and f_E, firms for which the expected profits is high enough to at least cover their sunk cost of entry "survive" the cost draw and start producing, while those for which the expected profits are less than f_E exit the product market. This defines the free entry condition of the model as

$$\int_0^{c_D^l} \pi_D^l(c)dG(c) + \int_0^{c_X^l} \pi_X^l(c)dG(c) = f_E \tag{4}$$

where the right- hand side is the expected profits of producing in location l.

In picking the optimal quantity and price combination for supplying locally, each firm in l takes as given the number of varieties produced locally, N^l, and those produced in the other location, N^h. It also takes as given the respective average prices, $\bar{p}^l$ and $\bar{p}^h$, charged in both locations. This is a case of monopolistic competition whereby profit maximization in each firm's pricing and production choices leads to equilibrium prices and quantities that can be expressed in terms of cost thresholds as

$$p_D^l(c) = \frac{1}{2}\left(c_D^l + c\right) \tag{5}$$

and

$$p_X^l(c) = \frac{\tau^h}{2}\left(c_X^l + c\right) \tag{6}$$

where $p_D^l(c)$ and $p_X^l(c)$ are the local and "export" components of the price $p^l(c) = p_D^l(c) + p_X^l(c)$ charged by producers in location l for their locally sold and exported quantities of

$$q_D^l(c) = \frac{L^l}{2\gamma}\left(c_D^l + c\right) \tag{7}$$

and

$$q_X^l(c) = \frac{L^h}{2\gamma}\left(c_X^l + c\right) \tag{8}$$

respectively where $q^l(c) = q_D^l(c) + q_X^l(c)$, and where L^l and L^h are the respective sizes of the product markets in the two locations, measured in terms of the aggregate number of consumers in each location.

Equations (5) through to (8) lead to $\pi_D^l(c)=\frac{L^l}{4\gamma}\left(c_D^l+c\right)^2$ and $\pi_X^l(c)=\frac{L^h\tau^h}{4\gamma}\left(c_X^l+c\right)^2$ as the expressions for the maximized profits from local sales and from exports, respectively of equation (4)—the free entry condition. Assuming a specific functional form for $G(c)$ in that equation leads to relatively precise predictions about the effects of market size, transport costs, and product substitutability as demand side determinants of aggregate productivity. Thus if $G(c)$ is a Pareto distribution with shape parameter k, such that

$$G(c)=\left(\frac{c}{c_M}\right)^k, \qquad c\in\left[0, c_M\right] \tag{9}$$

and we assume that the differentiated product is produced in both locations the free entry condition (4) reduces to

$$L^l(c_D^l)^{k+2}+L^h\rho^h\left(c_D^h\right)^{k+2}=\gamma\phi f_E \tag{10}$$

where $\phi=2(k+1)(k+2)(c_M)^{k+2}$, $1/c_M$ is the technological lower bound of productivity, k is higher the more concentrated is the industry in the sense of the number of high cost potential producers being higher relative to that of all potential producers; and $\rho^l=\left(\tau^l\right)^{-k}\in\left(0,1\right)$ is a parameter monotonically decreasing in transport costs. The parameter ϕ is increasing in the maximum cost threshold c_M and therefore decreasing in the technological lower bound, $1/c_M$. It also increases in the shape parameter of the cost distribution, k. Indeed ϕ is increasing in the variance (or dispersion) of the cost distribution of equation (9) and therefore in the dispersion of productivity across firms.

Equation (10) can be solved for the upper cost bound, c_D^l, of local supply as

$$c_D^l=\left[\frac{\gamma\phi f_E}{L^l\left(1+\rho\right)}\right]^{\frac{1}{(k+2)}} \tag{11}$$

on the simplifying assumption that transport costs are symmetric between the two locations, so that $\rho^l=\rho^h=\rho$.

D. Demand side and supply side determinants of aggregate industry productivity: Predictions set 1

Equation (11) relates the upper bound of the firm level distribution of the unit cost of production, c_D^l, of the industry's product to demand side and supply side factors, that is, to (γ, ρ, L^l), on one hand, and ϕ and f_E, on the other . But the model's assumption of a single factor-of production and constant- returns- to- scale technology of production means that every one of its predictions about the upper bound of the unit cost of production, c_D^l, readily translates into one about the lower bound of average factor productivity, $1/c_D^l$. The equation indeed contains the full range of the predictions of the model about the effects of product market demand factors on industry level aggregate (or average) productivity.

The predictions are that, other things being equal, average productivity, which is given by $1/c_D^l$ is higher: 1) the larger is the product market (i.e., the larger is L^l); 2) the greater is the ease with which consumers substitute between varieties in that market (i.e., the smaller is γ); and 3) the smaller is the cost of transport/shipment of varieties to and from other locations (i.e., the greater is ρ).

These three predictions should be set against a fourth one, also read from equation (11) via equation (10). That is, given all three demand side factors, industry wide average (or aggregate) productivity, $1/c_D^l$, is higher the smaller is the non-recoverable cost of entry, f_E, which in turn depends on a host of supply side factors, such as legal regulation of entry.

Equation (11) also implies that aggregate productivity, $1/c_D^l$, is decreasing in a second set of supply side factors, namely, ϕ and k, of which the first measures the variance of the cost distribution given by equation (9) and the productivity distribution underling that cost distribution. As a positive correlate of the variance of the cost distribution given by equation (9) and that of the underlying productivity distribution ϕ also inversely measures the concentration of the population of firms relative to that of the "marginal firm,"— that is, to the least productive or highest cost firm. It is also the case that ϕ is higher the higher is k and the greater is c_M. Average productivity, $1/c_D^l$, is therefore higher the lower c_M and the smaller is k.

E. Market size as a factor in firm dynamics, selection and productivity: Predictions set 2

Given that $c_D^l = p_{Max}^l$ in equilibrium, we can express the quantity of locally sold output in equation (5) in terms of p_{Max}^l as well as in terms of the minimum cost threshold. For the same reason, we can replace p_{Max}^l by c_D^l in equation (3) to solve for the number of firms selling in location l as

$$N^l = \left(\frac{2(k+1)\gamma}{\eta}\right)\left(\frac{\alpha - c_D^l}{c_D^l}\right) \tag{12}$$

This shows that, the number of firms operating in either of the two location of activity is higher: 1) the higher is average productivity, $(=1/c_D^l)$ in that location; 2) the lower is substitutability across varieties (or the higher is γ); and 3) the higher is overall aggregate demand for the industry's product in the sense of α being higher or η being smaller. At the same time, the model indicates that neither market size nor transport costs have any bearing on the number of firms operating in the economy or on its distribution across locations.

This results because the number of firms selling produce in location l, N^l is composed of two parts. One part consists of firms that are producing in l. The other part comprises those that are producing in the other location, h, and shipping at least some of it to l. The first of these is given by the product $G(c_D^l)N_E^l$, where N_E^l is the number of entrants to the industry in location l -- that is, firms that have incurred the sunk entry cost f_E, to make a cost draw from that location, but of which only a fraction given by $G(c_D^l)N_E^l$ do eventually decide to produce while the balance exit the industry having discovered that their cost draw is too high. The second group of sellers in location l, that is, exporters from the other location is likewise given by $G(c_X^h)N_E^h$, where N_E^h, the number of entrants to the industry in location h so that $G(c_X^h)N_E^h$ represents the fraction of them that not only survive to produce in h but are productive enough to cover the cost of shipments to sell in l. The number of firms selling in l can therefore be expressed as the sum:

$$N^l = G(c_D^l)N_E^l + G(c_X^h)N_E^h \tag{13}$$

This can be solved for the rate of entry in either location by making use of the assumption of symmetric transport costs between the two locations used in equation (11), to get the number of entrant in l as

$$N_E^l = \frac{(c_M)^k}{1-\rho^2}\left[\frac{N^l}{(c_D^l)^k} - \rho\frac{N^h}{(c_D^h)^k}\right]. \tag{14}$$

But only some of the N_E^l entrants "survive" in the sense of drawing at or below the cost threshold c_D^l. The number of such survivors (or actual producers) is given by $N_D^l = G(c_D^l)N_E^l = \left(\frac{c_D^l}{c_M}\right)^k N_E^l$, which can be restated by using (11) as

$$N_D^l = \frac{1}{1-\rho}\left[N^l - \rho N^h\left(\frac{c_D^l}{c_D^h}\right)^k\right] \tag{15}$$

Equations (12) to (15) lead to further predictions about the relationship between the number of firms, the scale of entry, the survival of firms, and three exogenous variables, namely: market size, transport costs and decline in such costs (or increases in connectivity between locations).

F. Market size and number of sellers: More firms sell in larger markets

The most immediate of these predictions it that more firms will sell in the larger of the two markets. This is in the sense that if $L^l > L^h$ then $N^l > N^h$ and vice versa. This follows from equation (12) via equation (11), which implies that $c_D^l < c_D^h$, which means that the larger market does not tolerate less productive firms as much as the smaller market.

G. Market size and the scale of entry: larger markets attract more entrants

The larger market also attracts more entrants in the sense that $N_E^l > N_E^h$ if and only if $L^l > L^h$. This follows from the fact that, by per equation (14), the scale of entry itself increases with the number of sellers already in the market, and is higher where productivity is higher. Because the larger market has more sellers and more productive firms, it attracts more entrants than the smaller market.

H. Market size and the scale of production: more is produced in larger markets

Not only does the larger market attract more entrants, it also has more survivors and thus supports more production than the smaller market in the sense that more of the entrants in the larger market produce for the local market than there are local suppliers in the smaller market. This is in the sense that $N_D^l > N_D^h$ if $L^l > L^h$. This follow from equation (15) since the larger market has more sellers -that is, $N^l > N^h$- and has more productive local producers, that is, $c_D^l < c_D^h$

III. Estimation Strategy & Results

Although presented in terms of the simplest case of production with a single-factor a single input the Melitz-Ottaviano model readily applies to the general setting of production with multiple inputs. In that setting the appropriate measure of productivity corresponding to $1/c_D^l$ in equation (11) as the lower bound of productivity defining the exit threshold in firm turnover is that of the lowest value of total factor productivity (TFP).

In this study we measure a plant's efficiency by its physical productivity (TFPQ) which is defined as:

$$TFPQ_i = \frac{q_i}{x_i} = \frac{\omega_i x_i}{x_i} = \omega_i \tag{16}$$

where q_i is the physical quantity of goods products by plant i, x_i is the value of inputs used to produce q_i, and ω_i represents the plant's "true" level of technical efficiency.

Most empirical studies which examine the link between selection mechanism and increased aggregate productivity are not based on physical productivity.[3] Instead, they are based on revenue productivity (TFPR) which is defined as

$$TFPR_i = \frac{p_i q_i}{x_i} = p_i \omega_i \tag{17}$$

where p_i is the price charge for its product by firm i. While these two measures are highly correlated, they are not identical as a plant's technical efficiency, ω, is only one factor that determines its profitability.

Measurement issues arise when producers within the same industry charge different prices due to variation in idiosyncratic demand or market power. There are two likely sources of within-industry price dispersion in industries where firms produce homogeneous goods. First, demand variation across local markets might arise due to transport costs. In such markets, plants derive market power due to horizontal price differentiation and high demand producers are more likely to survive (and set higher prices), even if they are less efficient than their low-demand rivals. Second, long-run supplier-buyer ties might exist between established plants and their consumers. In this case, consumers do not view all suppliers as identical, even if they produce goods that are physically identical. Of course, this case is identical to that of spatial demand variation as long as such relationships are based on horizontal rather than vertical product differentiation. In both cases, "high" productivity firms (in terms of TFPR) may not be technically efficient (in terms of TFPQ).

To confound matters, the two measures of productivity are inversely correlated with price (FHS, 2008). TFPR is positively correlated with price while TFPQ is negatively correlated. In our sample, we observe these correlations. We estimate the correlation between ln(TFPR) and ln(Price) to be 0.03 while that for ln(TFPQ) and ln(Price) is -0.60. To untangle these effects, we construct measures of TFPR and TFPQ for all plants (with 10+ employees) which produce flour, white pan bread, and cinder blocks in Ethiopia. Two criteria were used to choose these products. First, consumers are likely to view these products as identical in terms of their physical attributes. For

[3] The exception, of course, is the study by Foster, Haltiwanger, and Syverson (2008).

example, if two trucks showed up at a location with a flatbed full of cinder blocks, consumers would be indifferent as to which supplier they used. Second, we chose only industries where the number of observations was large enough to be statistically meaningful. To generate consistent measures of quantities across different industries, each product is measured in terms of its weight in kilograms.

Table 1 presents the descriptive statistics of these industries. As can be seen in the table, large productivity dispersion exists within each industry. For example, a plant at the 90^{th} percentile (in terms of TFPR) of its industry distribution produces between 1.9 and 2.8 times more than a plant at the 10^{th} percentile *with the same level of inputs*. These estimates are well within the range of those estimated for other countries. Syverson (2011), for example, estimates the ratio of 90^{th} to 10^{th} percentiles of TFPR at 1.92 for 4-digit industries in the United States. Similar to other studies, we find that the estimated within-industry productivity dispersion is larger when it is estimated by TFPQ rather than TFPR.

Finally, Table 1 reports the average (annual) entry rates and exit rates for each industry. We define a plant's entry date as the year when it was established (as reported by its owner or manager). On average, entry rates vary between 2% and 9% per year. Similarly, we define a plant's exit date as the last year that it was observed in the census. On average, exit rates vary between 15% and 21%. We should point out that substantial effort was made to ensure that our exit variable reflects a "true" exit from the market rather than a situation in which a plant was not interviewed during a given year. In the panel, there are cases when some firms are not interviewed in one or two years and then reappear in later years. These firms are not included in our estimated exit rates. To ensure that this is the case, we drop the last two years of the census (2010 and 2011) to check that firms which exited in 2009 do not reappear in either 2010 or 2011. In our sample, there are no cases where a firm is not observed for more than two years and then reappears. Our estimated exit rates are similar to those which have been reported for other manufacturing industries in Africa.

A well-known empirical finding is that firms in larger cities are more productive than those in smaller cities (see Combes and Gobillon, 2015 for a recent survey of this literature). While this "stylized fact" is based almost entirely on data from rich countries, it is likely that many of the same mechanisms which generate higher productivity in larger cities (e.g., agglomeration economies and competition-driven selection) operate in low-income settings as well. To date, however, few empirical studies examine these mechanisms for developing countries. This study is the first to our knowledge which measures the productivity gap (in terms of either TFPR or TFPQ) between African cities which vary in size. As can be seen in Figure 2, there is no evidence that the TFPR distribution for plants in Addis Ababa (Ethiopia's primate city) lies to the right of the TFPR distribution for plants in Ethiopia's secondary cities. However, the right-hand tail of the TFPR distribution in Addis Ababa is much longer than that in secondary cities, indicating that Addis contains a higher proportion of the country's most productive plants (in terms of productivity).

To investigate the potential impact of selection mechanisms on aggregate (industry) productivity, we begin by estimating a set of demand equations for each product. At the plant-level, differences in price and demand within a market reflect the strength of producers' horizontal demand differentials. To measure these effects, we estimate the following demand equation:

$$lnq_i = \beta_0 + \beta_1 lnp_i + \sum \delta_i Year_i + \lambda Income_{mt} + \eta_{it} \tag{18}$$

where q_i is the physical output of plant i in year t, p_i is the price of plant i in year t, $Income_{mt}$ is the average income (measured by the luminosity of night lights)[4] in plant's local market m, and η_i is a plant-year specific disturbance term. Estimating equation (3) using OLS could lead to biased estimates of the price elasticity β because plants respond to demand shocks in η_{it} by raising prices. Following the identification strategy proposed by FHS (2008), we use plant-level TFPQ as an instrument for producers' prices. Given that our measure of TFPQ reflects producers' idiosyncratic technologies (ω_i), it should be correlated with prices but orthogonal to idiosyncratic

[4] See the appendix for a description of how we use the night lights data to measures the level of demand in a plant's local market.

demand. Indeed, plant-level TFPQ explains 58% of the variation in producers' prices, after controlling for plant and year fixed effects.

Table 2 reports the result of estimating the demand isoelastic curves separately for each product. As expected, the IV estimates are more elastic than the OLS estimates which suggests the presence of simultaneity bias in the OLS estimates and TFPQ is an appropriate instrument for price. All (IV) estimated price elasticities exceed one in absolute value. This result is consistent with the hypothesis that producers have market power and are operating on the elastic portion of their demand curve which is what we would expect in markets characterized by horizontal product differentiation. In addition, we use the results from the IV demand equation to estimate idiosyncratic demand shocks. To do this, we continue to follow FHS (2008) by using the plant-level residual from the IV demand equation which we then add wwwto the level of local income (as measured by the luminosity of night lights).

Next, we examine how the characteristics of entering and exiting firms differ in terms of our four key variables: TFPR, TFPQ, price, and idiosyncratic demand shocks. These differences are computed by regressing each these variables on the entry and exit dummies as well as product-year fixed effects. We estimate both unweighted and weighted OLS regressions where the weights are producer-level revenues. The results of this exercise are reported in Table 3. We focus only on the results from the weighted regression analysis. It is clear from Table 3 that exiting firms have both lower TFPR and prices compared to incumbent firms. While the difference in mean TFPQ between firms is negative (as expected), it is not significant. Interestingly, we find that entrants charge significantly higher prices than incumbents and are subject to fewer idiosyncratic demand shocks. These results contradict those found for the US by FHS (2008) where entrants charge lower prices and, as a result, are an important driver of aggregate (industry) productivity growth.

To investigate whether entry and exit patterns vary by city size, we re-estimate the weighted OLS regressions but now split the sample by location—that is, we examine the difference in means separately for plants which are located in Addis Ababa and secondary cities. Table 4 reports these results. Similar to the full sample, exiting plants in Addis Ababa have lower TFPR and prices than incumbent firms. This pattern does not hold for plants located in secondary cities. While the coefficients on TFPR and TFPQ for exiting firms are negative, they are not significant. Exiting plants, however, in both Addis Ababa and secondary cities have lower prices than incumbent firms. These patterns are not repeated for entrants. In Addis Ababa, the difference in means between entrants and incumbents are not significantly different for any of our four, key variables. Entrants in secondary cities enter the market with higher prices than incumbents and are subject to fewer idiosyncratic demand shocks. These results provide some evidence that selection mechanisms are stronger in Addis Ababa than in secondary cities.

Finally, we get to main issue of the paper: whether selection mechanism in African cities are driven by productivity (TFPQ) or profitability (TFPR). We do this in two ways. First we estimate probit regressions where the dependent variable takes the value of one if the firm exits the market in that year. The explanatory variables in this model are our four key variables: TFPR, TFPQ, price, and idiosyncratic demand. The results of this regression analysis are reported in Table 5. We then estimate another specification of this model where we include as additional controls the plants' age, capital stock, transportation costs, location, and level of local competition in the market where it operates. Location is defined as a dummy variable which takes the value of one if the plant is located in Addis Ababa. Local competition is defined as producer density—the number of producers per kilometer in the market where the plant is located. A plant's market is defined as the urban agglomeration (as defined by night lights) where it is located. The results of these regressions are reported in Table 6.

When we estimate the parsimonious model (Table 5), we find similar results to those reported in Table 3. Firms with lower TFPR are more likely to exit the market than those with higher TFPR while the coefficient on TFPQ is negative but not significant. These results, however, are reversed when we add the additional controls (Table 6). In the full specification, TFPQ is both negative and significant, indicating that less efficient (higher cost) firms are more likely to exit the market than more efficient (lower cost) firms. This finding is consistent with the view that selection mechanisms in urban areas are associated with increased factor reallocation within industries. Interestingly, the coefficient on TFPQ becomes significant only when we include producers' transport costs in the specification. Notice that the coefficient on transport costs is negative, indicating that firms with lower transport costs are more likely to exit the market. One interpretation of this result is that firms that operate in markets with high transport costs are shield to some extent from competition, potentially slowing down the industrialization process.

IV. Concluding Remarks

In this paper, we investigate the potential impact of selection mechanisms in raising plant-level productivity in Ethiopia. Specifically, we are interested in how firms' entry and exit contribute to the pace of factor reallocation and TFP growth within industries—and whether these processes are accelerated in larger cities. We carry out this analysis using data from the Ethiopian census on manufacturing firms which covers the period 2000 to 2010. Importantly, these data include information on plants' physical outputs and their prices which allow us to distinguish between revenue-based measures of total factor productivity (TFPR) and those based on physical productivity (TFPQ). Our analysis reveals that these two measures generate very different results, suggesting that physical productivity measures (TFPQ) are better suited to examining firm dynamics when local producers have some degree of market power. In addition, we find evidence that less efficient (higher cost) firms are more likely to exit than their more efficient (lower cost) rivals—but only when we control for producers' transport costs. In urban areas, firms with lower transport costs are more likely to exit, suggesting that such firms are shielded to some extent from competition, potentially slowing down the process of industrialization.

References

Baily, M., Hulten, C. and D. Campbell. 1992. “Productivity Dynamics in Manufacturing Plants,” *Brookings Papers on Economic Activity: Microeconomics*, 1887-267.

Bairoch, P. 1988. *Cities and Economic Development: From the Dawn of History to the Present.* Translated by Christopher Braider Chicago: University of Chicago Press, 1988.

Combes, P.P. and L. Gobillon, 2015. “The Empirics of Agglomeration Effects.” In: Duranton, G., Henderson, V., and Strange, W. (Eds.), *Handbook of Regional and Urban Economics*, vol. 5A. (Amsterdam: North-Holland).

Fay, M. and C. Opal, 2000. “Urbanization without Growth: A not so Uncommon Phenomenon,” Policy Research Working Paper 2412, World Bank, Washington, DC.

Jedwab, R. and R.D. Osei. 2013. “Is Structural Transformation Different in Sub-Saharan Africa? New Historical Evidence from Ghana,” unpublished paper, George Washington University.

Foster, L., Haltiwanger, J., and C. Syverson. 2008. “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitabiliy?” *American Economic Review* 98(1) : 394-425.

Gollin, D. R. Jedwab, and D. Vollrath. 2016. “Urbanization with and without Industrialization,” *Journal of Economic Growth* 21(1): 35-70.

Hall, C. and C. Jones. 1999. “Why do Some Countries Produce So Much More Output per Worker than Others? *Quarterly Journal of Economics* 114(1): 83-116.

Hsieh, C-T. and P.J, Klenow. 2009. “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics* 124(4): 1403-1448.

Melitz, M.J. and G.I.P. Ottaviano. 2008. “Market Size, Trade and Productivity”, *Review of Economic Studies* 75: 295-316.

Prescott, M. E.C. 1998. “Needed: A Theory of Total Factor Productivity,” Journal of International Economics 89:1216-1233.

Restuccia, D. and R. Rogerson. 2008. “Policy Distortions and Aggregate Productivity with Heterogeneous Establishments,” *Review of Economic Dynamics* 11: 707-720.

Restuccia, D. and R. Rogerson. 2013. “Misallocation and Productivity,” *Review of Economic Dynamics* 16: 1-10.

Salop, S. 1979. “Monopolistic Competition with Outside Goods,” *Bell Journal of Economics* 10(1): 141-156.

Syverson, C. 2004a. “Market Structure and Productivity: A Concrete Example,” *Journal of Political Economy* 112(6) : 1181-1222.

Syverson, C. 2004b. “Product Substitutability and Productivity Dispersion,” *Review of Economics and Statistics* 86(2): 534-550.

Syverson, C. 2011. “What Determines Productivity?” *Journal of Economic Literature* 49(2): 326-365.

United Nations, Department of Economic and Social Affairs, Population Division. 2015 World Population Prospects: The 2015 Revision, New York.

Venables, A. 2016. “Breaking into Tradables: Urban Form and Urban Function in a Developing City,” unpublished paper, Department of Economics, University of Oxford.

World Bank. 2016. *Breaking Down Barriers: Unlocking Africa’s Potential through Vigorous Competition Policy*. Washington, DC: World Bank Group.

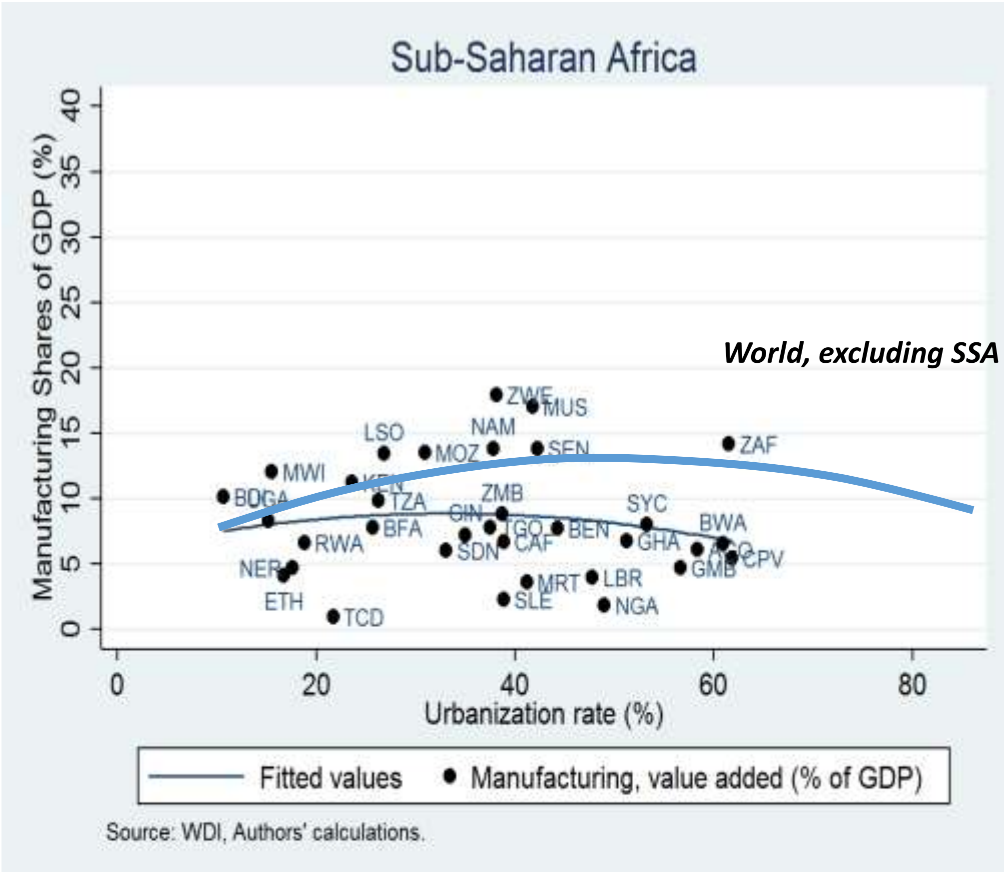


Figure 1: Urbanization and Economic Development
Source: Authors' calculations based World Bank Economic Surveys, 2015.

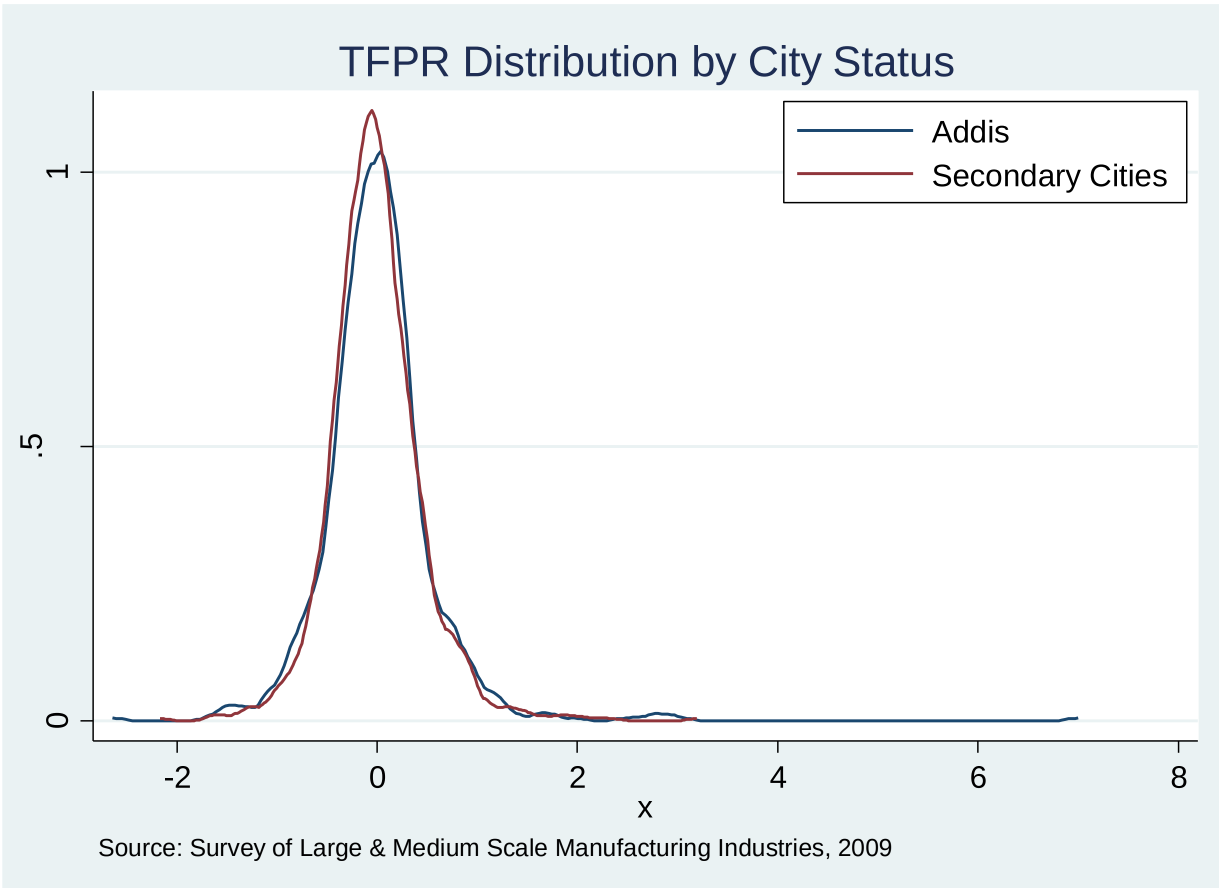


Figure 2: The Distribution of TFPR by 4-Digit Industry in Ethiopia

Table 1: Summary Statistics			
	Flour	**Pan Bread**	**Cinder blocks**
Entry	0.09 [0.46] (855)	0.02 [0.34] (786)	0.09 [0.49] (1,380)
Exit	0.15 [0.36] (855)	0.21 [0.41] (794)	0.28 [0.48] (1,362)
Firm size	44.39 [81.62] (858)	44.88 [293.18] (777)	21.42 [52.21] (1,258)
Ln(Capital)	12.93 [1.46] (810)	10.28 [2.71] (774)	13.04 [1.64] (1,246)
Ln(Transport Costs)	10.38 [1.96] (727)	9.04 [2.07] (582)	8.15 [2.11] (903)
Addis Ababa	0.24 [0.43] (860)	0.52 [0.50] (794)	0.27 [0.45] (1,362)
90^{th}-10^{th} percentile in TFPR	1.88 (803)	2.22 (757)	2.83 (1,144)
90^{th}-10^{th} percentile in TFPQ	2.73 (805)	22.92 (760)	3.42 (1,154)
90^{th}-10^{th} percentile in Ln(Price)	2.58 (853)	3.83 (793)	3.75 (1,353)

Notes: Standard deviations reported in brackets; number of observations reported in parentheses.

Table 2: Estimated Price Elasticities by Product				
	IV Estimates		OLS Estimates	
	Price coefficient (β)	Income Coefficient (λ)	Price coefficient (β)	Income Coefficient (λ)
Flour	-2.250***	0.411*	-0.734***	0.316*
	(0.35)	(0.22)	(0.09)	(0.19)
	[615]	[615]	[656]	[656]
Bread	-1.872***	0.313	-0.751***	0.130
	(0.17)	(0.35)	(0.06)	(0.25)
	[666]	[666]	[693]	[693]
Cinder blocks	-1.715***	-0.267	-0.827***	-0.061
	(0.23)	(0.17)	(0.06)	(0.14)
	[900]	[900]	[1,078]	[1,078]

Notes: Results from estimating demand isoelastic curves separately for each product. Standard errors clustered by plant are shown in parentheses. Average income is proxied by average luminosity of night lights in the urban agglomeration where the plant is located. *** significance at 1% level, ** significance at 5% level, * significance at 10* level.

Table 3: Evolution of Key Variables

Full Sample	Unweighted Regression		Weighted Regression	
	Exit	Entry	Exit	Entry
TFPR	-0.0928***	-0.0294	**-0.0831*** **	-0.0400
	(0.028)	(0.026)	**(0.028)**	(0.026)
TFPQ	-0.130**	-0.0228	-0.0959	-0.0164
	(0.066)	(0.061)	(0.067)	(0.065)
Price	-0.1520***	0.1420***	**-0.1510*** **	**0.1310*** **
	(0.041)	(0.043)	**(0.043)**	**(0.046)**
Idiosyncratic demand shocks	0.1190	-0.3290***	0.119	**-0.321*** **
	(0.108)	(0.085)	(0.112)	**(0.089)**

Note: This table presents the difference in means between entering & exiting plants. These differences are computed by regressing each of our four key variables (lntfpr, lntfpq, price, and idiosyncratic demand) on entry and exit dummies as well as product-year fixed effects. Standard errors clustered at the plant level are shown in parentheses. Our pooled sample has between 2,200 and 2,400 observations depending upon the specification. Weighted regressions are weighted by producer-level revenues. *** and ** indicate significance at 1% and 5% level, respectively.

Table 4: Evolution of Key Variables by City Type				
By City Type:	Addis Ababa		Secondary Cities	
	Exit	Entry	Exit	Entry
TFPR	**-0.1430**** **(0.057)**	-0.0151 (0.048)	-0.0429 (0.030)	-0.0553 (0.032)
TFPQ	-0.0580 (0.136)	0.1710 (0.125)	-0.1030 (0.076)	-0.0647 (0.075)
Price	**-0.2690***** **(0.086)**	-0.0996 (0.101)	**-0.104**** **(0.048)**	**0.208***** **(0.047)**
Idiosyncratic demand shocks	0.1690 (0.219)	0.1980 (0.185)	0.100 (0.128)	**-0.382***** **(0.101)**

Note: This table presents the difference in means between entering & exiting plants. These differences are computed by regressing each of our four key variables (lntfpr, lntfpq, price, and idiosyncratic demand) on entry and exit dummies as well as product-year fixed effects. Standard errors clustered at the plant level are shown in parentheses. Our pooled sample has between 2,200 and 2,400 observations depending upon the specification. All regressions are weighted by producer-level revenues. *** and ** indicate significance at 1% and 5% level, respectively.

Table 5: Selection on Profitability or Productivity

	Exit	Exit	Exit	Exit
TFPR	-0.224^{**}			
	(0.0783)			
TFPQ		-0.0503		
		(0.0295)		
Price			-0.153^{***}	
			(0.0279)	
Demand				-0.0153
				(0.0158)
Observations	2271	2282	2566	2427

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Selection on Profitability or Productivity with added Controls

	Exit	Exit	Exit	Exit
TFPR	-0.115			
	(0.104)			
Addis	0.0682	0.0557	0.0697	0.107
	(0.0907)	(0.0894)	(0.0879)	(0.0878)
Producer density	0.0850	0.0918	0.0847	0.0701
	(0.0655)	(0.0648)	(0.0641)	(0.0638)
Firm age	-0.170**	-0.166**	-0.174**	-0.160**
	(0.0563)	(0.0555)	(0.0546)	(0.0563)
Capital	-0.0387	-0.0369	-0.0341	-0.0299
	(0.0216)	(0.0212)	(0.0210)	(0.0209)
Transport costs	-0.0958***	-0.100***	-0.100***	-0.0946***
	(0.0228)	(0.0216)	(0.0215)	(0.0220)
TFPQ		-0.0747*		
		(0.0341)		
Price			-0.158	
			(0.0857)	
Demand				-0.00605
				(0.0219)
Observations	1661	1671	1726	1648

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

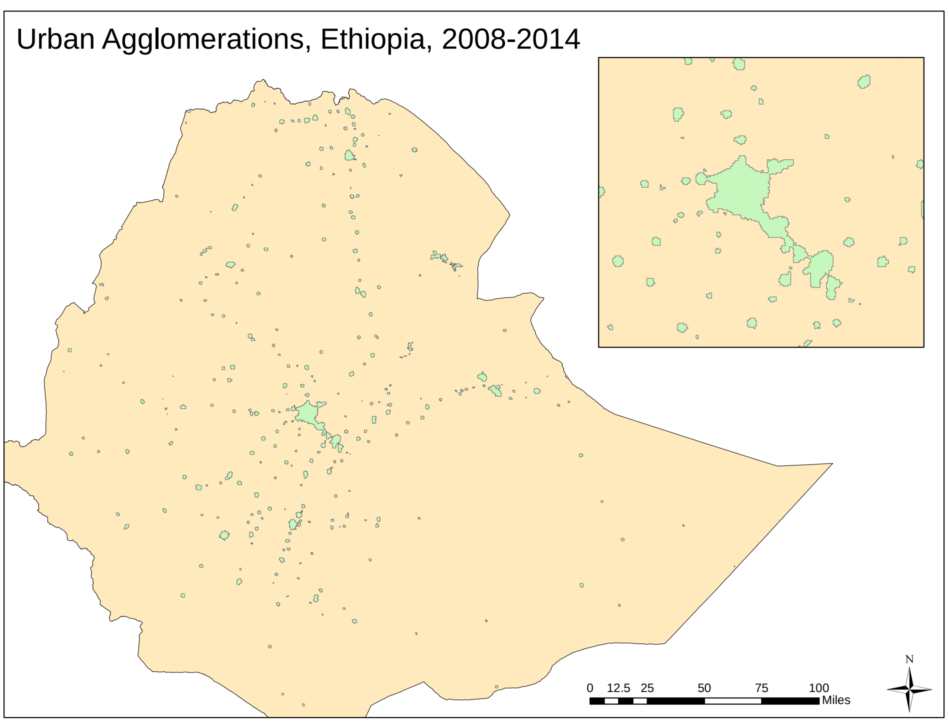
Appendix I: Night Lights Data

The Night Light (NTL) data that are used in this research are based on a study by Nigmatulina (2015). In this study NTL data are used to determine city boundaries or "urban agglomerations". The dataset that is created contains information on population, luminosity and area size of urban agglomerations in 38 countries in Sub-Saharan Africa. With ArcGIS, the urban agglomeration boundaries in this study are achieved by intersecting lit areas with city points that are at least populated once on the Citypopulation.de website. A lit area is an area that is lit at least twice between 2008 and 2012. An urban agglomeration is thus created when a unit is lit at least twice and the area contains at least on city point. To deal with urban agglomerations that cross borders, areas are separated along borders to become separate entities. When there are more populated points within a lit area in a given year, the population for this area then is based on a sum of the points.

Within the urban agglomerations luminosity per year, for all years between 1992-2012, is calculated as the sum of values for each pixel inside the unions. The study recognized that the censor doesn't record values below 5 properly and therefore all values between zero and five are treated equal, and all positive values that are less than 6 as equal to 6. In addition, the study explains that gas flared areas are correctly identified for all years and have been excluded from the calculation.

We have extracted from this data set the Ethiopian urban agglomeration boundaries to create a variable for economic activity per year per urban area. Our first step was to clip the urban agglomeration data with the country boundary of Ethiopia. Map 1 shows us the spatial distribution of urban areas across Ethiopia and the difference in size between the Addis Ababa agglomeration and the rest of the urban areas.

Map1: Urban Agglomerations, Ethiopia, 2008-2014



Source: Nigmatulina (2015) Cities in Africa Project Documentation

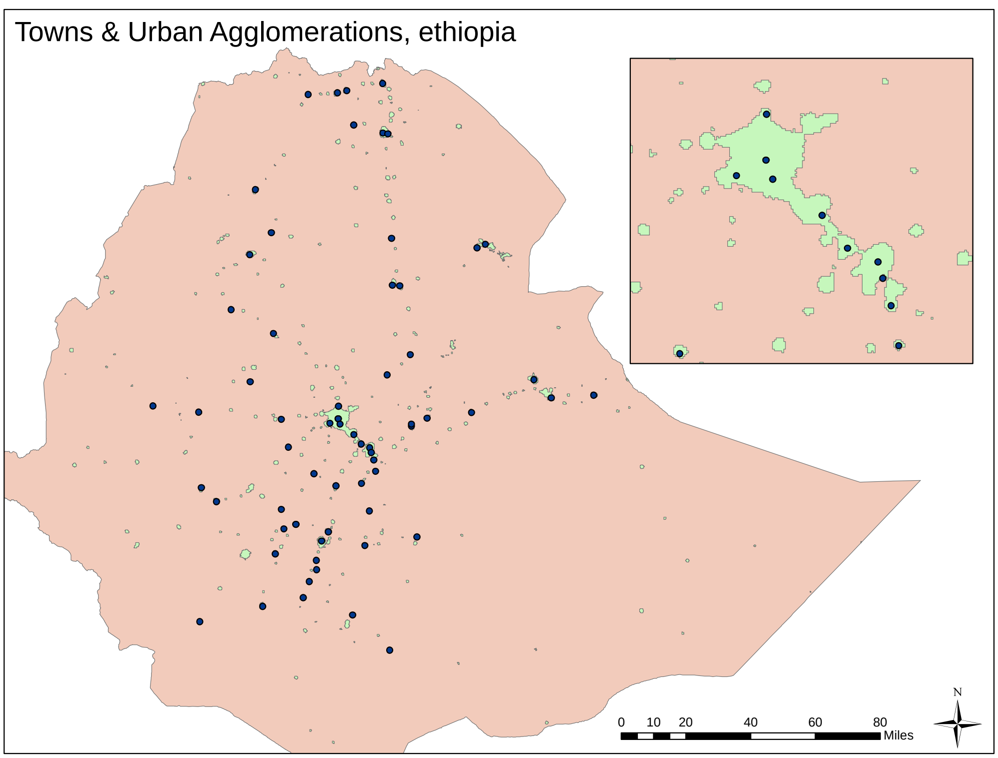
To match the urban agglomeration shape file with the panel data set, we georeferenced the towns of our data set. The coordinates of the towns from our panel data set are extracted and geocoded with GoogleMaps geocoder. The missing coordinates are substituted by a variety of online sources, such as openstreetmap. The coordinates of our towns are then added to the shape file and linked to the urban agglomerations by using a spatial join. Nearly all of our towns fall within one of the urban agglomerations. For one town, with one firm, we couldn't find the coordinates and we dropped this observation from the dataset. Two of our towns, both of them with one firm, didn't fall within an urban agglomeration and we dropped them from our dataset. Two of the observations that were named as towns in the dataset, containing 3 firms in total, were actually names of wereda's and therefore not matching any of the urban agglomerations and thus dropped from the dataset. The below table shows the towns and firms that were dropped form the dataset:

Table 1:

Town name	Number of firms	% of total dataset
Betterbecho	1	
Bebeka	1	
Deneba	1	
Akaki	2	
Cheliya	1	

Source:

Map2: Towns & Urban Agglomerations, Ethiopia



Source: Nigmatulina (2015) Cities in Africa Project Documentation